

# 基于云模型的 BP 算法改进

柴日发,徐文骞,曾文华

(杭州电子工业学院计算机分院,浙江 杭州 310037)

**摘要:**云模型是定性定量间的不确定转换模型,它将概念的模糊性和随机性集成在一起。文中提出一种利用云模型来有效避免 BP 算法陷入局部极小的方法,该方法通过基于云模型和输入参数区间划分的学习因子自适应调整算法来实现。该算法在复杂非线性分类(阴阳图)情况下进行了计算机仿真。

**关键词:**神经网络;算法;模型;自适应

**中图分类号:**TP183 **文献标识码:**B

## 1 云模型——定性和定量之间转换的有力工具

### 1.1 定义<sup>[1]</sup>

设  $X$  是一个普通集合  $X = \{x\}$ , 称为论域。关于论域  $X$  中的模糊集合  $A$ , 是指对于任意元素  $x$  都存在一个有稳定倾向的随机数  $\mu_A(x)$ , 叫做  $x$  对  $A$  的隶属度。如果论域中的元素是简单有序的, 则  $X$  可以看做是基础变量; 如果论域中的元素不是简单有序的, 而根据某个法则  $f$ , 可将  $X$  映射到另一个有序的论域  $X'$  中,  $X'$  中有一个且只有一个  $x$  和  $x$  对应, 则  $X'$  为基础变量, 隶属度在基础变量上的分布称为云。

云是用语言值表示的某个定性概念与其定量表示之间的不确定性转换模型, 云的数字特征用期望值  $Ex$ 、熵  $En$ 、超熵  $H$  三个数值表征。它把模糊性和随机性完全集合到一起, 构成定性和定量相互间的映射, 作为知识表示的基础。其中  $Ex$  是云的重心位置, 标定了相应的模糊概念的中心值。  $En$  是概念模糊度的度量, 它的大小反映了在论域中可被模糊概念接受的元素数, 即亦此亦彼性的裕度。  $H$  反映了云的熵的离散程度, 即熵的熵。云发生器通过输入三个数值特征就形成合乎条件的云滴, 从而将一个定性概念通过不确定性转换模型定量地表示出来<sup>[2]</sup>。例如“十几公里”就是一个不确定的语言值, 它的云模型表示如图 1 所示。

当给定三个数值特征和特定的  $x = x_0$ , 满足上述条件的云滴  $drop(x_0, y_i)$  的组合称为  $X$  条件云。当给定三个数值特征和特定的  $y = y_0$  时, 产生满足上述条件的云滴  $drop(x_i, y_0)$  的组合称为  $Y$  条件云。

### 1.2 定性规则的表达

#### (1) 一条定性规则的表达<sup>[3]</sup>

一条定性规则的形式化描述为:

If  $A$  then  $B$ ; 其中  $A, B$  为语言值表示的对象。

对照语言原子与云的对应关系, 可以方便地运用云对象来构造定性规则。图 2 说明了定性规则的构造方法 ( $CG$  为云发生器)。

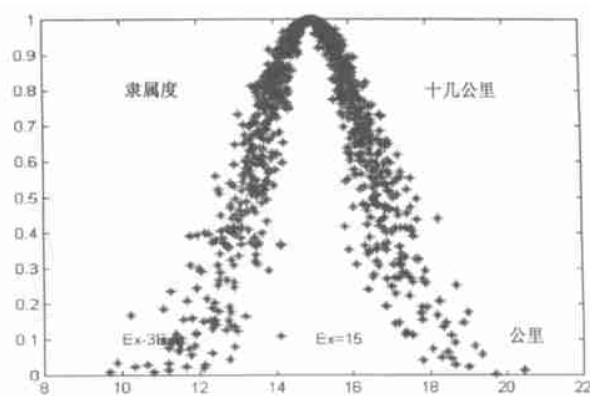


图 1 语言值“十几公里”的云模型表示

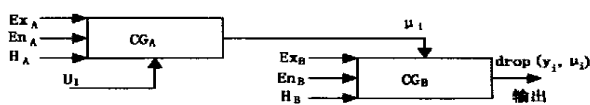


图 2 单规则生成器

我们用带  $X$  条件的云对象和带  $Y$  条件的云对象来构造单条定性规则。图 2 中,  $CG_A$  表示对应输入平面语言值  $A$  的带  $X$  条件的云对象,  $CG_B$  表示对应输出平面语言值  $B$  的带  $Y$  条件的云对象。当输入平面的某一特定的输入值  $U_1$  刺激  $CG_A$  时,  $CG_A$  随机地产生一组  $\mu_i$  值。这些值反映了对应定性规则的激活强度, 而这组  $\mu_i$  又控制输出平面的  $CG_B$  定量地产生一组随机云滴  $drop_i(y_i, \mu_i)$ 。从单规则激励机制示意图 3 中, 我们可以发现它存在固有的不确定性。对应一个固定的输入  $U_1$ , 首先  $CG_A$  产生了一组随机的输出值  $\mu_i$ , 将输入变量的不确定性传递到输出变量空间, 而  $CG_B$  在同一个  $\mu_i$  的控制下, 再次输出一组随机云滴  $drop_i(y_i, \mu_i)$  (参见图 3)。因为对应一固定的输入值, 输出空间中的  $y_i$  具有不确定性, 所以这一推理系统的实现对不确定性具有良好的继承性和传递性。可见, 云模型构造的定性规则较好地解决了不确定性的传播和更新问题。

收稿日期:2001-09-11

## (2) 多条定性规则的表示<sup>[2]</sup>

多条定性规则的形式化描述为:

$\text{If } A_1 \text{ then } B_1;$

$\text{If } A_2 \text{ then } B_2;$

$\text{If } A_n \text{ then } B_n.$

对照单规则生成器的思路,可以方便地运用它来构造多规则生成器。输入平面的某一特定的输入值  $U_i$  刺激各单规则生成器的前件  $CG_{Ai}$  ( $i = 1, 2, \dots, D$ ), 每一  $CG_{Ai}$  随机地产生一组  $\mu_{ij}$  ( $j = 1, 2, \dots, J$ ) 值。这些值反映了对应定性规则的激活强度,而多组  $\mu_{ij}$  又控制输出平面的  $CG_{Bi}$  ( $i = 1, 2, \dots, D$ ) 定量地产生多组随机云滴群  $\text{drop}_{ijk} (y_{ijk}, \mu_{ij})$  ( $k = 1, 2, \dots, K$ ; 各组云滴的数目—— $K$ 值的多少与激活强度  $\mu_{ij}$  的大小有关)。将所有的云滴输入到逆云发生器中(给定符合某一正态云分布规律的一组云滴作为样本  $(x_i, \mu_i)$ , 产生云所描述的定性概念的三个数字特征值  $(E_x, E_n, H)$ , 即从定量到定性的转换,其软件或硬件实现称为逆云发生器),生成的期望值可输出,或者取所有云滴中的随机值输出。

## 2 BP 算法改进

1985 年 Rumhart 等人提出的 BP 算法是目前神经网络学习算法中使用最为普遍的算法<sup>[4]</sup>, 由于 BP 算法学习慢且不易收敛,许多文献提出了一些改进方法,但存在一个问题:各种改进算法都仅解决问题的某一方面,各种算法的优势没有得到互补,从而不能有效解决局部极小和学习速度慢的问题。

BP 算法的学习过程是通过调整权重使期望值和神经网络输出值的均方差趋于最小实现的:

$$W_{ij}(t+1) = W_{ij}(t) + \frac{\partial E(t)}{\partial W_{ij}(t)}$$

其中  $W_{ij}(t+1)$ ,  $W_{ij}(t)$  分别为神经元在第  $t+1$  次和第  $t$  次调整后的权值系数;  $E(t)$  为均方差;  $\eta$  为学习速率。

为了加快网络的学习速度,通常是引入动量项。动量项使得调节向着底部的平均方向变化,不致产生大的摆动,即动量起到缓冲平滑的作用。带动量项的调节公式为:

$$W(t+1) = \frac{\partial E}{\partial W} + W(t)$$

其中  $W(t+1)$ ,  $W(t)$  分别是神经元在第  $t+1$  次和第  $t$  次的权重改变矩阵。为动量系数,一般取 0.9 左右<sup>[4]</sup>。

BP 算法的改进主要是通过对学习速率自适应调整来实现。BP 自适应调整算法<sup>[4]</sup>利用整个网络状态的知识自动地寻找一个合适的值,使得网络学习总是向误差曲面的底部平均方向变化。但是由于基于 BP 算法的网络误差曲面往往存在不少的局部最小点,因而使网络得不到最佳的结果。随机训练算法<sup>[4]</sup>利用随机过程,概率和能量关系来调节连接权,

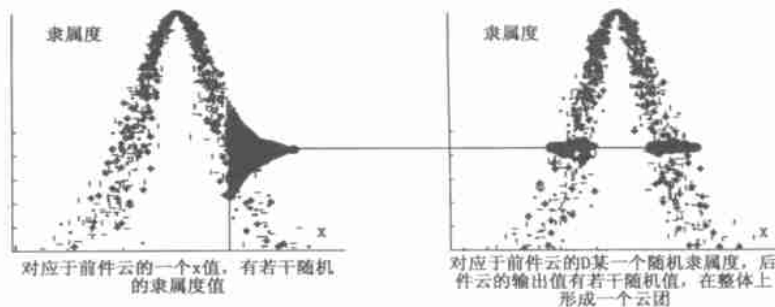


图3 单规则激励机制示意图

随机调节虽然可能使网络学习性能变差,但从总体上看,却可以保证网络获得最佳效果,即找到网络最小值。结合上述两类改进算法的优点,本文提出基于云模型和参数区间划分的改进算法。

对于学习速率的调整,就一般的经验来说就是当误差连续变小时,加大学习速率以加快收敛速度。当误差连续变大时,则减小学习速率以得到较小的调节。如果在调节过程中加入一定的随机过程则可以在收敛速度和全局最优的选择中得到一个折衷的方案。因而这种调整的一般规则为(假设  $e$  为误差值):

$\text{If } e < 0 \text{ and } \text{abs}(e) \text{ 较大 then 采用较大的学习速率和动量系数;}$

$\text{If } e < 0 \text{ and } \text{abs}(e) \text{ 较小 then 采用较小的学习速率和动量系数;}$

$\text{If } e > 0 \text{ and } \text{abs}(e) \text{ 较小 then 采用较小的学习速率和动量系数;}$

$\text{If } e > 0 \text{ and } \text{abs}(e) \text{ 较大 then 采用最小或零学习速率和动量系数。}$

云模型是定性定量间的不确定转换模型,它将概念的模糊性和随机性集成在一起,因此上述的规则可以用基于云模型的多规则生成器实现。我们以误差值或能表征误差方向的指标为输入参数来实现多规则生成器的  $X$  条件云,以反映误差调整倾向的学习速率云和动量系数云来实现多规则生成器的  $Y$  条件云。则这种云多规则生成器使随机过程和学习速率自适应调整有机地结合在一起而得到一较优的方案。云模型所蕴含的随机过程保证了神经网络在总体上能获得最佳效果,即找到网络最小值。而云模型多规则生成器的规则倾向性又保证了神经网络在实现上的收敛速度。

但是根据高斯函数的“3 $\sigma$ ”原理,当云发生器的输入参数在区间  $[E_x - 3E_n, E_x + 3E_n]$  之外时,  $\mu$  值趋于零。由于当输入参数超出多规则生成器的所有  $X$  条件云的上述区间时,所有的  $X$  条件云将输出几乎相等的趋于零的值,这时多规则生成器将起不到规则选择的功能,规则的调整倾向性失去作用,学习速率的调整完全处于随机状态。而此时的误差变化已经很大,学习速率的调整方向已明确,如果仍采用带随机过程的算法,将使网络不能收敛或极大地降低了网络性能。因

此为使网络得到最佳的效果,我们将输入参数区间划分为多规则生成器作用区和非作用区,当在非作用区时可以采用一般的学习速率自适应调整算法。但由于这时误差绝对值过大,而学习速率的调整倾向已明确,所以本算法中的所谓的一般的学习速率自适应调整算法也有特殊之处,即为当负误差绝对值过大时,用极大的学习速度和动量系数调整,当正误差绝对值过大时,用极小的学习速率和动量系数调整。因此本算法加入第五条规则:

if abs(s) 过大 then 采用一般的学习速率自适应调整算法。

至此,可以将本文提出的基于云模型的 BP 改进算法描述如下:

```
(1) flag = 0; // flag 用于标志算法的转换
(2) 得到输入参数 ui; // (2) ~ (4) 在循环体内
(3) if flag = 1 then 采用一般的学习速率自适应调整算法;
if 输入参数 ui 在多规则生成器作用区之内 then flag = 0;
(4) if flag = 0 then 采用云多规则生成器调整学习速率和动量系数;
if 输入参数 ui 在多规则生成器作用区之外 then flag = 1。
```

### 3 仿真

一般的神经网络对于线性可分的系统进行分类很容易实现,但对于复杂非线性可分的系统分类时就很难得到较好的结果。阴阳图如图 4 所示,阴阳图的阴阳两类相互交融,你中有我,我中有你,由于图中小圆的面积相对比例较小,所以阴阳图的分类极难实现,它可以作为算法改进是否成功的有效测试工具。



图 4 阴阳图



图 5 阴阳图的仿真结果

要用本文的改进算法,首先是多规则生成器的输入参数的实现。我们以平方误差梯度分量的符号来估计与最小均方误差的距离,以平方误差梯度分量的符号的正号连续出现的次数和负号连续出现的次数作为多规则生成器的输入参数。用一般的自适应调整算法运行 BP 网络得到输入参数的数据源,在除去奇异值后(即为误差绝对值过大的情况),采用逆云发生器得到 X 条件云的所有数字特征,以此构造多规则生成器的 X 条件云。多规则生成器的 Y 条件云根据已有的经验直接构造。本文仿真中构造的多规则生成器的数字特征如下:

前件 (- 3.5166, 0.4284, 0.1051) (- 1.3425, 0.2195, 0.0695)  
(1.5499, 0.4308, 0.1216) (6.8261, 1.0887, 0.3015)

后件 (1.4000, 0.0600, 0.0100) (1.1500, 0.0600, 0.0100)  
(0.8500, 0.0600, 0.0100) (0.6000, 0.0600, 0.0100)

上述多规则生成器的输出值是学习速率的增大或缩小的比例系数,即  $\text{rules}(\text{ipara})$ , 为学习速率, ipara 为输入参数。本文的多规则生成器的作用区间为  $[-6, 10]$ , 在这区间之外用一般的自适应调整算法实现。当负误差绝对值过大时 ( $\text{ipara} < -6$ ), 取比例系数为 1.8 来调整, 当正误差绝对值过大时 ( $\text{ipara} > 10$ ), 取比例系数为 0.2 来调整。以上输入参数的区间划分是对数据源的可视化后经分析得到的。

计算机仿真的结果表明上述做法是成功的, 仿真结果如图 5 所示。图中用了 20081 个数据点对用 10000 次训练周期训练好的网络进行了仿真, 结果基本上能够区分出两种不同的类。本网络用两层的双曲型 BP 网络和一层线性网络构成, 第一隐含层有 15 个神经元, 第二隐含层有 10 个神经元, 采用均匀采样的 1257 个数据点对此进行了训练, 且在仿真结果中以输出值的区间来分类, 即“阴” =  $[0, 0.5]$ , “阳” =  $[0.5, 1]$ , 这样在较大的误差水平下得到了较好的分类结果。

一般的 BP 改进算法也可以得到如图 5 所示的结果, 但是需要用 15000 次训练周期才能得到相近的结果, 而且从学习速率曲线的变化来看, 两者还是有很大不同, 本文算法的学习速率曲线的变化很明显, 体现了算法的随机性和规则的倾向性。两者的学习速率曲线如图 6 所示。(为得到清晰的曲线图, 这里仅用了 2000 次的训练周期, 但实际上用 2000 次还无法得到图 5 的效果)。

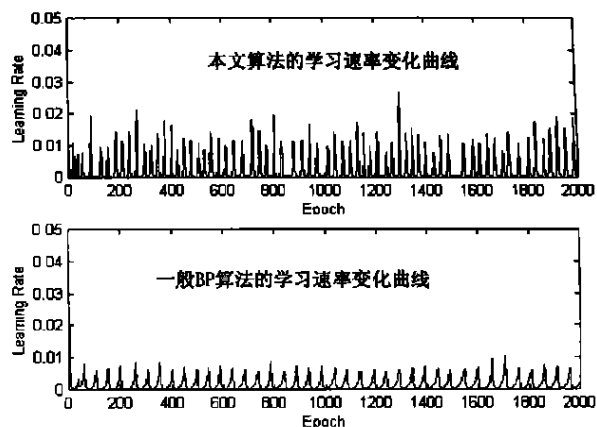


图 6 学习速率比较图

### 4 结束语

BP 算法有多种多样的改进, 但是对算法组合的研究还是比较少见。对于算法的组合最为关键的是对于算法转换标志的研究。本文采用输入参数区间划分来组合一般的自适应调整算法和多规则生成器算法是对算法组合研究的一种尝试。我们希望这种做法可以取各种算法的长处, 得到满意的结果。

- [1] 李德毅, 孟海军, 史雪梅. 隶属云和隶属云发生器[J]. 计算机研究和发展, 1995, 32(6): 15 ~ 20.
- [2] Li D Y. Uncertainty reasoning based on cloud models in controllers [J]. Journal of Computer Science and Mathematics with Application. H. SVIER SCIENCE, 1998, 35(3): 99 ~ 123.
- [3] 陈晖, 李德毅, 沈程智, 张飞舟. 云模型在倒立摆控制中的应用[J]. 计算机研究与发展, 1999, 36(10): 1180 ~ 1187.
- [4] 袁曾任. 人工神经网络及其应用[M]. 清华大学出版社, 2000, 78 ~ 104.



**柴日发**(1978.1 - ), 男(汉族), 浙江省江山市人, 2000年6月毕业于浙江大学经济学院, 现为杭州电子工业学院计算机应用技术专业硕士生。研究方向为人工智能及应用。

**徐文骞**(1976-),男(汉族),江苏镇江人,杭州电子科技大学研究生,研究方向为智能控制,网络数据库等。

**曾文华** (1964. 3 - ) 男 (汉族), 江苏省兴化市人, 1989 年 6 月毕业于浙江大学工业自动化专业获工学博士学位, 现为杭州电子工业学院计算机分院教授。研究领域为工业过程模型化及优化控制, 计算机先进控制, 智能控制。

( School of Computer Science , Hangzhou Institute of Electronics Engineering , Zhejiang Hangzhou 310037 ,China )

**ABSTRACT:** Clouds model is one for transformation between qualitative and quantitative knowledge ,whose representation reflects fuzziness and randomness contained in linguistic concepts. In this paper ,a clouds model is applied to BP algorithm improvement ,which avoids BP ANN trapping to local least value. This learning rate self - tuning method is based on clouds model and input parameter domain partition. Finally ,this algorithm is simulated under complex non - linear classification circumstances.

**KEY WORDS :** Nerual Network ;Algorithm ;Model ;Self - tuning

[illegible]

\* \* \* \* \*

COMPUTER SIMULATION (Bimonthly)

Started in Dec ,1984

|             |  |
|-------------|--|
| 主 管 单 位     | 中国航天工业总公司  |
| 主 办 单 位     | 中国航天机电集团公司第十七研究所   |
| 编 辑 出 版     | 《计算机仿真》杂志社   |
| 印 刷 装 订     | 中国航天科技集团公司第十九所照排印刷中心   |
| 主 编         | 吴连伟  |
| 编辑部主任       | 束志业  |
| 副 主 任       | 曾莲芝 邵光震  |
| 地 址         | 北京海淀区阜成路 14 号  |
| 邮 编         | 100037   |
| 电 话         | (010) 68767186   |
| 传 真         | (010) 68373574   |
| E - m a i l | <a href="mailto:kwcoltd@public.bra.net.cn">kwcoltd@public.bra.net.cn</a>   |
| 网 址         | <a href="http://www.kwsc.cncom.net">www.kwsc.cncom.net</a> <a href="http://www.compusimu.com">www.compusimu.com</a>  |
| 网 络 版 址     | <a href="http://www.chinajournal.net/jsjz">www.chinajournal.net/jsjz</a><br><a href="http://www.chinainfo.gov.cn/periodical">www.chinainfo.gov.cn/periodical</a> |
| 刊 号         | ISSN 1006 - 9348/ CN11 - 3724/ TP  |
| 邮 发 代 号     | 82 —773 (国内)<br>4651M (国外)   |
| 国 外 发 行     | 中国国际图书贸易总公司  |
| 广告许可证       | 京海工商广字第 0277 号   |
| 定 价         | 8.50/本 51.00/ 全年   |
| 期刊基本参数      | :CN11 - 3724/ TP *1984 *b *A4 *110 *zh *¥8   |

Responded by China Aerospace Corporation  
Sponsored by the 17th Institute of CASC  
Edited & Published by The magazine agency of COMPUTER  
SIMULATIONB  
Editor in Chief : Wu Lianwei  
Add: No. 14 Fucheng Road Haidian District Beijing China  
Zip : 100037  
Tel : (8610) 68767186  
Fax : (8610) 68373574  
E- mail : [kwcoltd@public.bta.net.cn](mailto:kwcoltd@public.bta.net.cn)  
URL : <http://www.kwsic.cncom.net> <http://www.compusimu.com>  
[www.chinajournal.net/jsjz](http://www.chinajournal.net/jsjz)  
[www.chinainfo.gov.cn/periodical](http://www.chinainfo.gov.cn/periodical)  
Journal Code : ISSN 1006 - 9348/ CN11 - 3724/ TP  
International Postal : 4651M  
Subscribed by China International Book Trading Corp  
P. O. Box. 399 Beijing China 100044